DOCUMENT RESUME

ED 070 524 PS 006 217

AUTHOR Arima, James K.; Gray, Francis D.
TITLE An Information Analysis of 2-, 3-, and 4-Word Verbal Discrimination Learning.
PUB DATE 72
NOTE 14p.; Paper presented at the Annual Convention of the American Psychological Association (80th, Honolulu, Hawaii, Spetember 2-8, 1972)

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS *Data Analysis; *Discrimination Learning; *Information Theory; *Task Performance; *Verbal Learning

ABSTRACT

Information theory was used to qualify the difficulty of verbal discrimination (VD) learning tasks and to measure VD performance. Words for VD items were selected with high background frequency and equal a priori probabilities of being selected as a first response. Three VD lists containing only 2-, 3-, or 4-word items were created and equated for information content at approximately 20 bits. Lists were presented for 6 trials at 1.5 or 3.0 sec. per item with 20 Ss per list. Only correct responses were reinforced. The uncertainty (H) in the response distribution was calculated for each item on each trial and summed for each list. The H remaining after Trial 6 averaged over all 3-sec. conditions was 8.4 bits. For the 1.5-sec. conditions, it was 15.2 bits, indicating that information was being processed at a constant rate. The amount of information processed per trial was also constant for each list condition, although differences among lists reached borderline significance (p .05) at the 3-sec. rate. It appeared quite feasible to apply information theory to VD learning as an absolute measure of task difficulty and performance. (Author)

FILMED FROM BEST AVAILABLE COPY

ED 070524

INFORMATION ANALYSIS OF 2-, 3-, AND 4-WORD VERBAL DISCRIMINATION LEARNING

JAMES K. ARIMA and FRANCIS D. GRAY
Naval Postgraduate School, Monterey, California

Information theory was used to quantify the difficulty of verbal discrimination (VD) learning tasks and measure VD performance.

Method. Words for VD items were selected with high background frequency and equal a priori probabilities of being selected as a first response. Three VD lists containing only 2-, 3-, or 4-word items were created and equated for information content at approximately 20 bits. Lists were presented for 6 trials at 1.5 or 3.0 sec. per item with 20 *S*s per list. Only correct responses were reinforced.

Results. The uncertainty (H) in the response distribution was calculated for each item on each trial and summed for each list. The H remaining after Trial 6 averaged over all 3-sec. conditions was 8.4 bits. For the 1.5-sec. conditions, it was 15.2 bits, indicating that information was being processed at a constant rate. The amount of information processed per trial was also constant for each list condition, although differences among lists reached borderline significance ($p < .05$) at the 3-sec. rate.

Conclusions. It appeared quite feasible to apply information theory to VD learning as an absolute measure of task difficulty and performance.

DIVISION 3

PS 006 217

U. S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

(Paper presented at the 80th Annual Convention of the American Psychological Association, Honolulu, 1972)

ED 070524

An Information Analysis of 2-, 3-, and 4-Word Verbal Discrimination Learning

James K. Arima and Francis D. Gray

Naval Postgraduate School

PS 006217

Underwood and Freund (1969) have shown that learning difficulty varies as the number of choices in a verbal discrimination (VD) item are varied when the total number of items in a list are held constant. Zacks (1969) has demonstrated that the total learning time tends to be invariant over various conditions of practice for a fixed task load. These characteristics of VD learning tasks would seem to have much in common with the information analysis of communication tasks (Garner, 1962). That is, the number of possible alternatives in any VD item could be readily quantified in terms of the amount of information contained in the item if the expected relative frequency (a priori probability) of choice of each word could be identified. The total amount of information in a list would then be the sum of the information contained in each item of the list. Similarly, the invariance of total learning time could be expressed quantitatively in information measures as the rate of information processing.

The information analysis of VD items with more than two alternatives might also provide additional information over conventional methods of analyzing VD learning. Conventional methods quite often use the percent of correct responses for analysis. This measure varies with the number of alternatives in an item, in addition to other aspects of the learning situation. Information measures, on the other hand, could provide an

absolute measure of learning that is independent of the number of alternatives in a VD item. Thus, the purpose of this research was to show that verbal discrimination learning could be quantified and analyzed using information theory.

Method

Three VD lists were constructed differing on the basis of having 2-, 3-, or 4-word items (figure 1). Within each item, the words were

- -

Insert figure 1 about here

- -

high in background frequency, had an equal a priori chance of being selected on the first trial, and were from similar or different categories for half the items. The determination of equal a priori probabilities was based on background frequency of the words as suggested by the frequency theory of VD learning of Ekstrand and Underwood (1966) and confirmed by exploratory testing. Background frequency was controlled by using only AA and A words in the Thorndike and Lorge (1944) norms. The words were selected from the category norms of Battig and Montague (1969). Words were selected for similarity from the same category and for dissimilarity from different categories. Since the choice of any alternative was equally likely, the information content of any item was the logarithm to the base 2 of the number of choices. Accordingly, a 2-word item had one bit of information; a 3-word item, 1.585 bits; and a 4-word item, two bits. The 2-, 3-, and 4-word lists were then equated for total information content at approximately 20 bits by having 20, 12, and 10 items per list.

Twenty different Naval Postgraduate School students were assigned to each list and given six presentations of the list. Only correct answers were reinforced, thus providing one bit of information to the subject regardless of the number of words in the item. The lists were presented at 1.5 secs. per item to half the Ss in each list treatment and at 3.0 secs. to the other half. Interitem times were also 1.5 and 3.0 secs. Thus, item length and presentation rate were factorially combined and similarity was a nested variable.

Results

The percent of correct responses are shown in table 1 and figure 2.

- - - - - - - - - - - - - - - - - - - -

Insert table 1 and figure 2 about here

- - - - - - - - - - - - - - - - - - - -

Analysis of variance showed lists, presentation rate, and trials to be significant at the .01 level and similarity significant at the .05 level. The rate by trials interaction was significant at the .01 level, and list by similarity interactions was significant at the .05 level. All findings were in the expected direction; similarity, however, was facilitative in only the 3- and 4-word lists.

The information analysis was conducted by first determining the distribution of choices for each item on each trial. From these distributions, the average uncertainty for each item, considered as a message set, was calculated. The item uncertainties were then summed and subtracted from the original 20 bits to provide the amount of information yet to be learned or transferred per trial. These uncertainties are shown in table 2 and figure 3. A rank chi-square analysis of variance showed item length to be significant at the .05 level for the 3.0 sec.

condition but not for the 1.5 sec. condition. The uncertainty remaining at trial six, averaged over the 1.5 sec. treatments was 15.2 bits; the abount remaining for the 3.0 sec. treatments was 8.4 bits.

- -

Insert table 2 and figure 3 about here

- -

Discussion and Conclusions

An analysis of the "percent correct" responses showed significant differences for item length, rate of presentation, and similarity. Nevertheless, the information measure could be used to produce approximately equal information transfer among the lists and an approximately constant rate of information transfer for two presentation rates differing by a factor of two. The borderline significance of item length in the 3.0 sec. treatment suggests that a uniform presentation rate per item for all item lengths may have resulted in a more difficult information processing task for items with more words.

It appears that information theory can be used to quantify the difficulty of a VD learning task. This capability, when fully developed and generalized, could have many applications in both applied and research settings by providing an absolute measure of the difficulty of a learning or training task or the information processing abilities of individuals.

References

Battig, W. F. & Montague, W. E. Category norms for verbal items in 56 categories. Journal of Experimental Psychology, 1969, 80, Monograph, No. 3, Part 2, whole.

Ekstrand, B. R., Wallace, W. P., & Underwood, B. J. A frequency theory of verbal-discrimination learning. Psychological Review, 1966, 73, 566-578.

Garner, W. R. Uncertainty and structure as psychological concepts. New York: Wiley, 1962.

Thorndike, E. L. & Lorge, I. The teacher's word book of 30,000 words. New York: Columbia University Press, 1944.

Underwood, B. J., & Freund, J. S. Verbal-discrimination learning with varying numbers of right and wrong terms. American Journal of Psychology, 1969, 82, 198-202.

Zacks, R. T. Invariance of total learning time under different conditions of practice. Journal of Experimental Psychology, 1969, 82, 441-447.

PS 006217

Table 1

Percent of Correct Responses per Trial

by Item Length and Presentation Rate

Trial	4-Words		3-Words		2-Words	
	1.5-Sec.	3-Sec.	1.5-Sec.	3-Sec.	1.5-Sec.	3-Sec.
1	29.0	30.0	33.3	34.1	48.5	46.5
2	28.0	41.0	43.3	46.7	54.0	64.5
3	32.0	42.0	41.7	65.0	55.0	76.0
4	34.0	55.0	55.0	65.0	64.5	79.0
5	42.0	62.0	50.8	73.3	66.0	83.5
6	40.0	69.0	60.0	82.5	76.0	87.5

Table 2

Uncertainty Remaining per Trial by Item Length and Presentation Rate

Trial	4-Words		3-Words		2-Words	
	1.5-Sec.	3-Sec.	1.5-Sec.	3-Sec.	1.5-Sec.	3-Sec.
1	16.9	17.0	18.2	17.5	18.2	18.4
2	18.0	17.8	16.9	16.3	18.6	17.4
3	17.3	16.9	16.5	13.3	18.4	13.0
4	17.1	14.1	15.6	13.2	15.9	12.6
5	17.1	13.5	15.6	11.5	16.0	10.7
6	16.6	10.3	15.0	7.1	14.0	8.2

Figure Captions

Fig. 1. Word lists for 2, 3-, and 4-word treatments. The first word in each item was used as the correct response. Dissimilar items are denoted with an asterisk.

Fig. 2. Percent correct responses by treatment and presentation time.

Fig. 3. Uncertainty remaining per trial for each treatment and presentation time.

Two-Word Treatments

murder	wine*	juice	doll*
apple	river*	book	lake*
iron	yard*	nail	swim*
tea	coffee	temple	rock*
table	chair	cotton	salt*
mother	father	bus	gun*
cat	dog	water	door*
eye	head	car	train
foot	mile	red	blue
corn	bean	hour	minute

Three-Word Treatment

door	temple	water*
brother	mother	father
green	blue	red
lake	book	table*
cotton	hours	salt*
apple	knife	cat*
iron	yard	doctor*
eye	head	foot
minute	hour	second
hill	river	rock
boat	train	car
nail	oil	swim*

Four-Word Treatment

mother	sister	brother	father
door	private	temple	water*
eye	foot	nose	head
swim	nail	wine	oil*
yard	doctor	iron	book*
yellow	blue	green	red
lake	rock	river	hill
cotton	salt	house	table*
cat	murder	knife	apple*
year	minute	hour	second

Fig 1

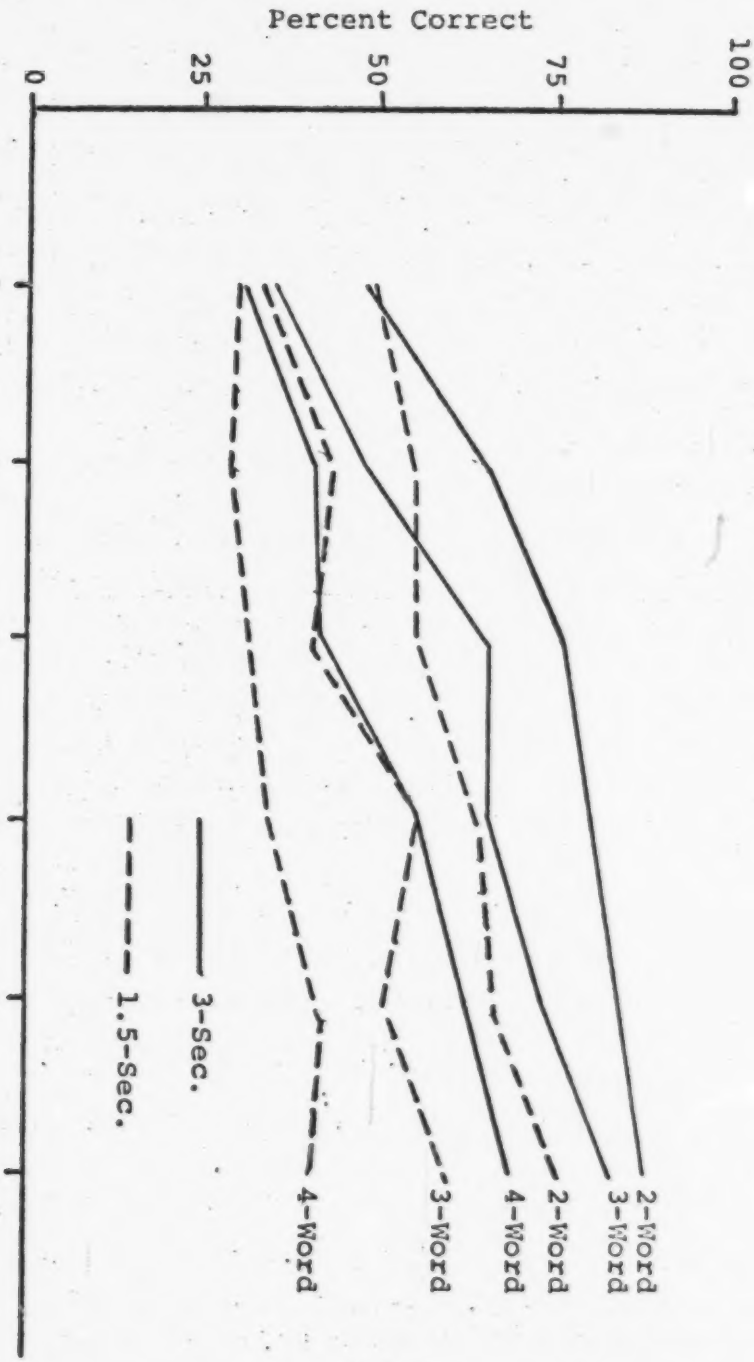

Percent Correct
0
25
50
75
100
1
2
3
4
5
6
Trials
2-Word
3-Word
2-Word
4-Word
3-Word
4-Word
3-Sec.
1.5-Sec.

Fig 2

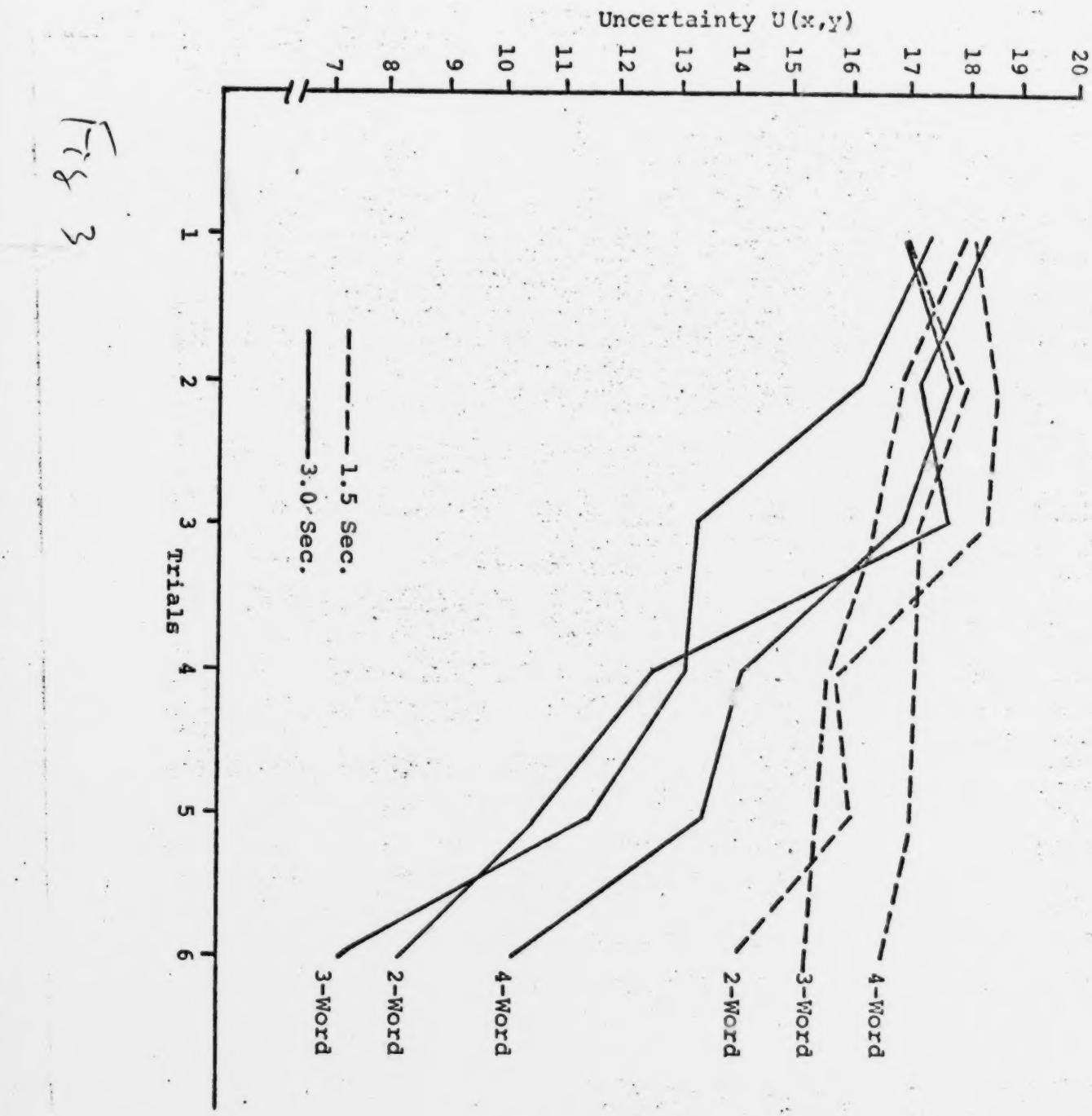

Uncertainty U(x,y)
20
19
18
17
16
15
14
13
12
11
10
9
8
7
1
2
3
4
5
6
Trials
1.5 Sec.
3.0 Sec.
3-Word
2-Word
4-Word
2-Word
3-Word
4-Word

Fig 3

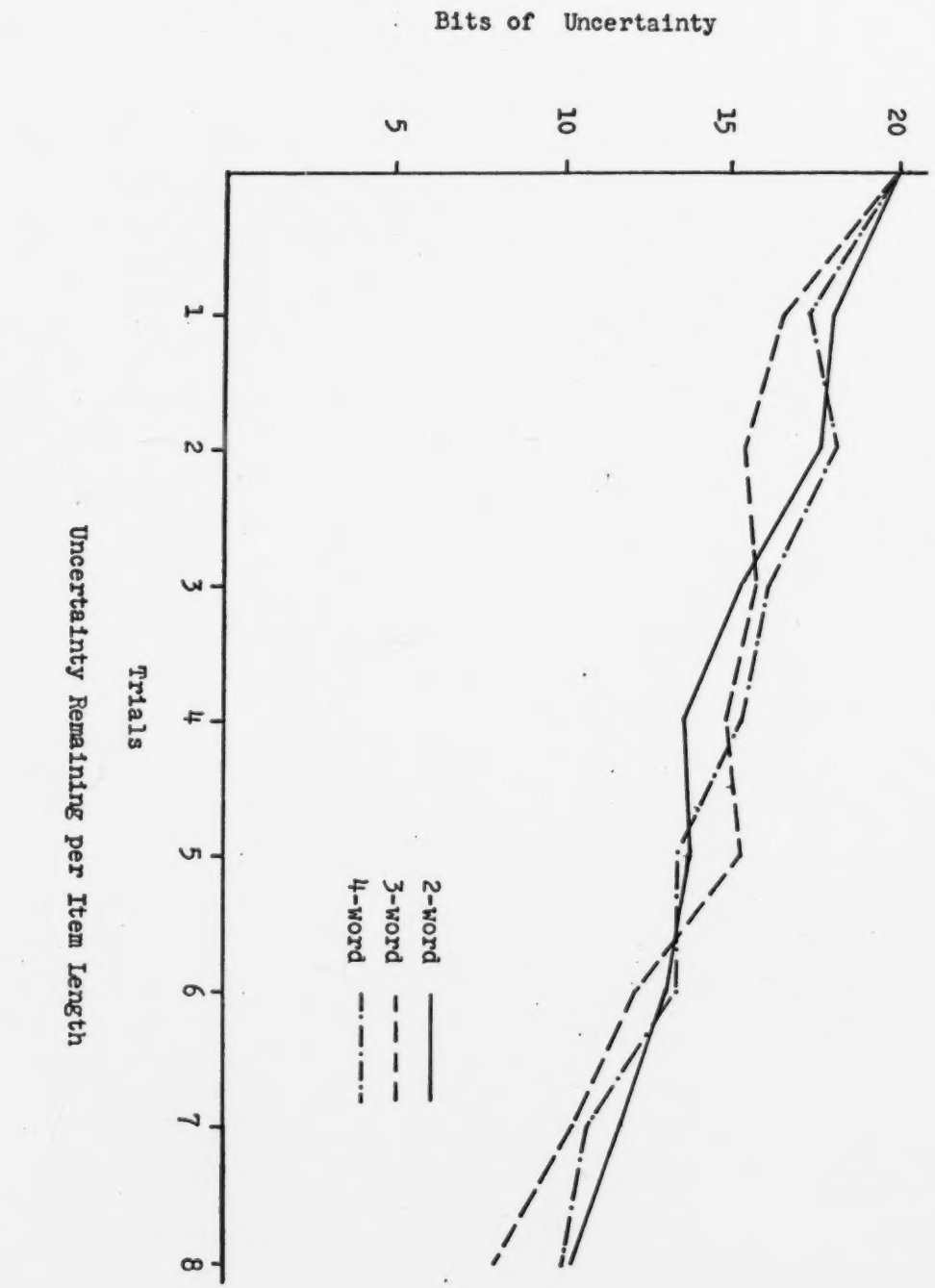
ADDENDUM

The study was redone with a constant presentation rate (two rates) over the item types as shown in the table below. The time intervals are in seconds. Each presentation had an item (stimulus) presentation time interval, as shown, followed by a blank interval of equal length. For example, for the 2-word item, the item was presented for 1-sec. followed by a 1-sec. blank interval, making the interstimulus interval two seconds and making the stimulus presentation rate ½-bit/sec.

The graph on the next page shows the uncertainty remaining after each of eight trials for the 2-, 3-, and 4-word conditions summed over the two presentation rates. There is no difference in the list or list by trials effects, and we have shown that we can take VD items of different lengths, made into lists of different lengths, and presented at different time intervals per item and have them all show the same information processing rate.

TIME INTERVALS FOR ITEM PRESENTATION

ITEM TYPE	PRESENTATION RATE	
	½ BITS/SEC	¼ BITS/SEC
2-WORD	1.0	2.0
3-WORD	1.6	3.2
4-WORD	2.0	4.0



Uncertainty Remaining per Item Length